PRE-SOLICITATION MEETING

FOR NIDDK CENTRAL REPOSITORY

PRERECORDED TAPE TRANSCRIPTION

November 28, 2001

Neuroscience Building
Room A1/A2
National Institutes of Health
Bethesda, Maryland

P R O C E E D I N G S

INTRODUCTIONS OF NIDDK STAFF:
Dana L. Harris, Program Analyst
Robert Hammond, Director, Division of Extramural Activities
James Everhart, Program Director
Sanford Garfield, Program Director
Stephen P. James, Deputy Director, Division of Digestive Diseases and Nutrition
Rebekah S. Rasooly, Program Director
Patrick Sullivan, Chief, Acquisitions Management Branch

DR. RASOOLY: The NIDDK is interested in developing a Central Repository that will carry out four tasks. I will review those tasks, which were outlined in the Request for Information. The first group of tasks is for archival storage of tissue samples, plasma samples, serum samples, cell lines, DNA samples collected in NIDDK-funded large, multi-center studies.

Now I did provide a table and, in fact, in the information I handed out there is a different table, which is an earlier version that estimates the number of samples that exist. That number is simply a collection.

It is a collection of everything that we have without regard to whether we are going to reposit it or not.

So one important piece of information I wanted to make clear here is that we really do not

have any intention of finding storage for nearly a million samples that are sitting in freezers right now. We are not planning to move all of them.

The “existing samples” column -- those numbers are simply some kind of working estimate of what we have and I would estimate that a very small fraction of those would actually be acquired by the storage facility, and that the primary purpose of the storage facility would be to house samples that are being collected now and in studies that will be completed in the future.

What this task involves many of you know better than I do but it involves storing the samples and labeling them in such a way that they are easy to retrieve in a low cost efficient manner with good back ups. One of our concerns, of course, with this is that for any sample that we have to acquire from another study that it may be difficult, and I would be interested in your thoughts on this, to acquire those samples and label them so that they could be retrieved efficiently and we could certainly talk about that some more.

The second group of tasks is completely different and these tasks relate to storage, long-term storage and, as it were, archiving of data from studies. We have run so many studies, which eventually do end or many of them end or we hope that

they will end. And when they end there are these large data sets that relate often to these biosamples and we are not exactly sure of what to do or how to maintain these data because it is not as simple as just keeping a CD in our office. It is a matter of people that can maintain the integrity of the data set, do an update if that is necessary, can search the data to find relevant samples. This is going to require people who can gain familiarity with different kinds of large data sets, enough to maintain and to search them and to help people who want to query those data sets in a logical way. And it does not exclude the possibility that the contractor might be interested in doing some data analyses, although that would not be specifically a part of this contract.

The third piece of the contract, the way we saw it, was the genetics. Supporting the genetics studies seems to be quite a different task from any of the others because this would involve serving as a real time repository, receiving blood samples, transforming cells so that you make immortalized cell cultures, and extracting the DNA and making those materials available either to the original researchers or to subsequent researchers, who are approved for access to these materials.

At this point we do not envision these tasks

involving any kind of genotyping or molecular analysis of the samples. At this point it does not seem like that is a routine enough task that could be accomplished efficiently by a contractor but one could envision that five or eight years from now that that would become a routine kind of task and something eventually that might be incorporated into the contract but we do not envision that for the first period of time.

And then, finally, and this is in some ways the most complicated aspect of what we want, we were hoping that we would be able to find contractors or groups of contractors that would be interested in serving as the real time repository for new studies. For example, if we start a new study next year that is investigating obesity or, you know, some endocrine disease or some aspect of diabetes, and it would be a large clinical study involving 10 or 15 or 20 sites, we would find contractors who would be willing to take on the task of being the dedicated processing facility for all samples from that study. That would be a mini-contract, a task as it were, that would be a long-term relationship with the investigators in the study. If another study came on board in another year, the contractor would take that on as the core lab for each of these studies.

This is a little bit more difficult kind of

task because it is hard for us to know exactly what we would expect the contractor to do. Each study has its own requirements in terms of the kinds of measurements, the samples they want to collect, and the measurements they want to carry out.

So this we thought that we would look towards having some kind of task order contract so that we could develop specific contracts or task for each study particularly and that this would be, you know, I think a great help to our investigators who are often ill-equipped to set up a core facility efficiently with the expertise to do these kinds of analyses.

So those are the four groups of tasks the way we saw them and we asked you, and I have gotten some feedback on this, for advice on how we might best approach finding a contractor or contractors to handle all these tasks. And so I thought what I would do is I would open the floor a little bit and ask if you have questions or comments or thoughts about how to organize all this.

DR. _____: I was wondering if you could issue a one single RFP and have these different tasks, the four different tasks that you list there, under that RFP and request offers to either submit a proposal for all or one of the tasks and that a contract or contracts could be awarded from that,

multiple contracts could, or a single contract could?

DR. RASOOLY: I am going to give a quick answer and then, Pat [Sullivan, Contracts Officer, NIDDK], I hope that you will correct me. Our feeling is that we have not decided exactly how we will do it, whether that is the best way to go, whether the best way to go is multiple RFPs, or just a single contract and let the contractor develop subcontracts for the different tasks and we have not made a decision exactly how to do that.

I think that it is extremely important and it certainly would be a technical evaluation criterion how the different pieces of the contract work together but I am not sure of the modality.

MR. SULLIVAN: I do not have anything to add to what Dr. Rasooly [NIDDK] said. This is something that we will need to review and analyze and take under advisement, and we will -- this is something we have not yet made a conclusion or a decision on.

DR. _____: (Not at microphone.) (Inaudible) of issuing multiple RFPs, that you could actually issue one RFP in which you could make multiple contracts of part of the statement of work or all of the statement of work to the contract?

DR. RASOOLY: That is not an uncommon practice.

DR. _____: Yes.

Mr. SULLIVAN: We have done that in the past.

DR. _____: Yes. I think it is important for all of us to start perhaps with the end in mind. Do you have -- let me be very bold and just say it very bluntly, do you have a number in mind in terms of what you are shooting for?

DR. RASOOLY: Dollars?

DR. _____: Yes.

DR. RASOOLY: No.

DR. _____: Okay.

DR. RASOOLY: In fact, that in some respects proves to be the most difficult part of all this planning because to some extent the cost drives the use of the repository. At the time that it becomes so expensive to carry out some aspect of repositing the samples, the value of actually doing that declines. So it is a very tricky and elusive kind of issue for us.

DR. _____: Excuse me. Is right now is one of the options -- I am not sure I heard this correctly at the beginning -- that for each of the four work areas that you summarized this morning to have at least one contract aimed specifically at those work areas? In other words, a unique contract per work area out of these three or four areas that were discussed this morning? Is one more likely to

be first out the door or still to be determined?

DR. RASOOLY: My plan is that we are going to try and address all these needs in one fell swoop, however we do it because once we are started and determined to do this, to do any one of the pieces and leave the others behind is not such a great idea. We certainly cannot archive the samples without archiving the data and for the ongoing studies we need to provide resources for them, especially given the large number of studies we are planning to launch in the future.

MR. SULLIVAN: What we have done traditionally in the past, for instance, for multi-center clinical trials where we have clinical sites and data coordinating centers and laboratories, we have made all those awards effective on the same date, and that is our thought on this one as well.

DR. _____: When I looked at the tasks what occurred to me was that A, C & D were clearly related to biomaterials but B had a mixture of data and biomaterials. And so I was suggesting that perhaps some of the tasks from B that dealt with biomaterials, for example, creating the lines, maintaining and storing cell lines and biosamples, I thought perhaps some of those could be moved from B up into A so that tasks A, C & D clearly deal with biomaterials only and B then is focused on data

issues only.

DR. RASOOLY: Right.

DR. _____: And that was a suggestion.

DR. RASOOLY: I think that is reasonable.

DR. _____: Okay.

DR. RASOOLY: I mean, the only issue, and I see this now with a study that I am working on right now, which is a nonprofit facility, but the only issue that I have seen is that, of course, every facility needs its own database. I mean, obviously no matter what you are storing you need a database. So we do need to provide, and the contractor needs funds and the contractor needs to provide the resources to show that they can actually track the sample from the minute it enters their facility, you know, in an ongoing way. And that data must integrate nicely with the associated data from the study that is being stored by the database contractor.

So to some extent we will, unless we are told otherwise, incorporate support of a database for each one of these components but data analysis and storage of associated data, I think, you are right should be exclusively with the data contractor.

DR. _____: I have got another question with the eventual use of the materials. Do you expect the specimens to be distributed to not only

NIDDK scientists but to any scientist that would possibly be studying and to what extent do you expect to distribute materials sometimes in the future?

DR. RASOOLY: Right. So the only point of making the repository is so that the samples will be available and certainly the samples will be of most interest to people studying the diseases that NIDDK traditionally studies. Who is most interested in a diabetic cohort, of course, is people studying diabetes.

On the other hand, I think that we will not restrict it to NIDDK researchers or even NIDDK-funded researchers. My feeling is that we will restrict it to people whose need for the samples is justified on the basis of the science and, of course, the ethical review and so on, given the fact that some of the samples will be in limited quantities.

DR. _____: As a follow up to that question, is it clear that all of these samples that were previously collected were consented for distribution in the manner that you suggest and, if so, do you see recertification of the samples in terms of their identity and qualities as an issue?

DR. RASOOLY: So this was the most fundamental issue that we began this process with and there was a decision from the institute director on down that no sample will go into this repository

unless the subject was specifically consented on having their sample in a repository. And so what that means in the case of studies that have been completed or that are ongoing is that the subjects will be reconsented, which again -- I mean that is the same issue of whether the samples can be reposited or not. If it is going to be impossible to reconsent the subjects then the samples are simply not accessible to the repository, but we are not willing to reinterpret consents at this point.

DR. _____: Whose job will that be, the reconsenting?

DR. RASOOLY: We have felt strongly – Jim [Everhart, NIDDK], do you want to speak to that?

DR. EVERHART: That clearly has to be done at the study sites that the patients were participating in. That is not going to be a function of the repository.

DR. _____: My concern is if the studies had been done several years earlier, it is very difficult to get patients to reconsent them. In genetic studies, in particular, if you are missing key people, the whole study is --

DR. EVERHART: This is why Dr. Rasooly said that only a very, very small portion of the already collected samples would actually be available because of these sorts of issues.

DR. RASOOLY: Our best success will be with studies which are in, what I would call from my microbiology background, the stationary phase now, where the patients are still being followed but are not being actively studies. They will be coming in for an exit interview and that is the most likely kind of study where we could get a reconsent in some kind of efficient way because they will be contacting the subjects anyway.

I would be eager to hear your thoughts. One thing that I felt is that the repository should not have any identifying information whatsoever associated with the samples. I have heard other points of view but I would be interested in any comments that you have on this.

It looks like there is broad agreement on this one. I mean, the fact that the repository or even the database repository has no identifying information prevents, for example, updating if one does a long term study ten years later to find out outcomes in terms of mortality or even, if it is possible, hospitalizations. If there is no identifying information that will not be possible so it is a concern that you will not be able to do very, very long-term follow-up of these subjects.

If I turn the question the other way, does anybody know of a repository where the identification

information is held?

DR. _____: The question is do you mean where it is possible to identify participants, individual participants in a study?

DR. RASOOLY: Right.

DR. _____: You mean that the links from the repository to the site that collected those samples will be broken.

DR. RASOOLY: That -- when I originally was thinking about this and I think when we were thinking about it that was the idea that the link between the alpha numeric identifier and the patient information would be broken. That is right.

DR. _____: And then with the transfer would come all of the medical data that was collected during the period that that person was in the study so that there would -- in other words, it would not be helpful to have just a freestanding sample where you did not know anything else about the participant that donated that sample?

DR. RASOOLY: Right. So the associated data would surely have the same alpha numeric identifier otherwise.

DR. _____: Have you considered using an escrow agent, that is an intermediate party who would keep data on subject and identity and study and provide an anonymized sample to the repository?

DR. RASOOLY: You know, I have actually never thought about that and I am not familiar with that. Is this a common device?

DR. _____: Yes, it is something that is done.

DR. RASOOLY: And these agents are typically employed by the contractor or not employed by the contractor?

DR. _____: Well, for an example, in studies that we do with HBDI, HBDI is essentially the escrow agent. They keep the information and the data and we get an anonymized sample. And if there are any requests for information, it goes to them and they subsequently issue a request for more samples or they request another sample and we do not know the connection.

DR. RASOOLY: Well, then that is, in effect, an option that I think has been outlined here.

DR. _____: Yes.

DR. RASOOLY: Is that whoever is doing the data analysis would simply, if that was another contract, tell the repository contract to pull these samples and the repository contractor knows nothing else except to pull those samples.

DR. _____: That is right.

DR. RASOOLY: So then I had -- was also curious as to the feeling of people in this group as

to the possibility of providing core services through a task order contract, whether that is practical and feasible to approach each study sort of as its own unit. And I did not know if people here with laboratory arrangements felt that this was a workable approach.

I guess the question being if the tasks are going to be different, is it possible for a contractor to acquire the kind of expertise that they would need to do to do a different set of blood measurements for one study than it would for another study? That is really what we are wondering or whether it is going to be de facto a new contract each time anyway so, you know, maybe not go this route at all. And that -- you know, without actually having run the business ourselves, we do not have a sense of how practical that is.

DR. EVERHART: We did see it as an advantage when we start one of these multi-center studies to be able to go to our existing contractor or contractors and say, you know, please tell us if you can do this and get a start on it rather than our current situation, which is sort of to have to go out and find groups to do the tasks that Dr. Rasooly outlined after we have already started a study. It becomes cumbersome, rushed, inefficient to do that. So this is the purpose but the question is because this --

from a business-side, does this seem to be a workable arrangement?

DR. _____: I think if you are -- if the tasks are centered around the products in this table then it is perfectly reasonable to expect one contractor to be able to deal with various individual investigators requesting different -- the storage of different products or processing products in a different way. I think that is perfectly reasonable.

DR. RASOOLY: Yes. Another question that people have disagreed about in my discussions with them is the process of acquiring samples from other sites. We have had a study. It has concluded. The samples are in two or three freezers in two or three different laboratories around the country. They are well marked and they have data associated with them.

And the question was, for example, do all those samples need to be completely relabeled if they are acquired by a repository for archival storage? What is involved in doing that? Do they have to be re-aliquoted into storage tubes? How flexible are repositories in acquiring samples from other places for archival storage?

We have no sense of that either. How adaptable are the systems? I see everybody smiling.

DR. EVERHART: I guess to put it another way, for those of you who may have done this or are

familiar with storing tissues, what are the issues involved that -- we talked about consent obviously but in terms of the actual samples, what are the -- what are the kind of quality control issues? What do you actually have to be able to do with those samples that you would be moving from one place to another?

DR. _____: We have run studies in the past where we have multiple types of studies coming into one repository. We do not have to re-label the samples necessarily as long as there are unique identifiers on them. We could keep those separate and just put them into one database.

DR. _____: Isn't there a requirement that the samples be barcoded and if the samples are not barcoded previously, wouldn't they have to be relabeled to meet that requirement?

DR. RASOOLY: We did write barcoding but we wrote it -- those words can leave as easily as they came within the requirements. The question is, isn't that state-of-the-art now that samples for ready retrieval are barcoded?

DR. _____: Yes, that is -- I mean, the way we currently do everything is we barcode samples and provide labels for samples in clinical trials but also, you know, if there is stuff out there from old studies that have not been barcoded, we could accommodate those either by relabeling or just use

them as they were with identifiers on them and hand type that into the information -- into the database and we could double data entry on those.

DR. GARFIELD (NIDDK): The way many clinical studies work is that the study group, in fact, through the coordinating center selects a laboratory that performs anywhere from five to 100 different kinds of analyses. And under (C) where you talk about new studies and the collection, I mean one through four really talks about the actual collection of the samples and maintaining them and working with the study. In five you talk about actually carrying out the laboratory processing.

And one thing that I have never been clear on is how that would work through a repository contract when the study group knowing what its study is really would like to, I would presume, have final say on who actually does the measurements on their samples and how would -- I mean, how would people here feel about how that part of this would actually be handled? You know, would you actually want to do that yourselves? Do you have the facilities to do that? Or would you, in fact, go through -- back through the study group to select who the best laboratory was?

DR. EVERHART: I think that is one of the uncertainties that was outlined about the task orders

is that each study is going to be unique, of course, in both their patient population and what is going to be tested. And I guess we were envisioning that once a protocol has been outlined and we more or less know what tests are going to be done and what samples are going to be obtained and when that it would be decided which study laboratories would be doing what and also offer that to the repository, this part of the contract. And if it is a test that only one or two labs in the country can do, then probably that is not going to be a repository. It is not going to be a function of this. But if it is something that this contractor can do or set up rather easily, and since they are going to be getting the samples anyway, then we would consider doing it.

I mean, that is an issue is that each sort of task for a new study is going to have its unique aspects.

DR. RASOOLY: One of the things that I am most concerned about, again referring to my own ongoing experience now with the study that I am working with, is integration between the pieces of the repository. So if we have a data contractor who is storing samples and we have a genetics contractor who is storing the blood samples from those patients, and we have, the archival folks who are storing -- is it feasible that those three entities will actually

work together and exchange data easily or is there something that we are missing here about ensuring that that -- that there is uniformity and easy ways for those people to communicate? Is there something that we should be building in here to simplify or to ensure that?

In other words, should -- I am thinking, for example, should one of the contractors have primary responsibility for labeling samples and determining what kind of labeling system will be used and that then be adopted by the others? Should there be a contractor working group? I am trying to think mechanistically how this might work out.

DR. _____: I think you highlight one of the problems associated with having a data repository for the entire contract. For example, in the cell repository end, one wants the data in a format that is immediately available for analysis that geneticists do, for example, and that is a very different kind of format and a very different kind of data management issue than, for example, the urine collections or the other collections. And I thought your statement presumed the fact that there would be a central data repository and I am not sure that is the best idea, although it is one way to go certainly.

DR. RASOOLY: So I think I should clarify

that. NIDA and NIMH and several other institutes have contracts whereby the data, the primary data from the study that are associated with the genetic samples are deposited in a database that is contract funded by those institutes. That is not what we are envisioning here at all.

What we are envisioning is that the data collected on the subjects will be collected by the study's data coordinating center. They are the ones that are -- after all, they are the researchers. They are the ones that are most familiar. They will collect the data. They will put it into a format. If it is a genetic study obviously there will be genetic data and so on. And that will be a complete data set. You know, I sort of see this in my own mind as CDs. You know, that will be that CD for that study.

And the data contractor will work with the data coordinating center to receive that CD, to receive that database, to understand enough how to work with it, what its structure is, what its underlying structure is, and so on so that they could search it but that getting the data into shape will be 100 percent local to the study that is generating the data and that is what we are envisioning. Otherwise it becomes unworkable.

The studies are extremely different and we

could not, I think, in any way hope to make it a single database, which brings me back to my question again of how you work integration among the different pieces, the people that are holding the CD, the people who are holding the tissue sample, and the people that are holding the DNA and the cell line?

It sounds like something we will have to do, huh? Okay. Yes, go ahead.

DR. _____: We have solved that problem at Duke with our collection centers. Our resource centers actually can generate their own data set number, acquisition number. From that we assign it, at Duke, a double coded number and then from that everything falls into place and how we sequentially send out those samples for further extraction, storage, ship them out to other collaborators. So that in way we also keep patient records, patient identity, basic research data separate from our data collection and data management center and we handle samples about the number that you envision for your repository center.

DR. _____: So this is from numerous different studies, different --

DR. _____: Absolutely. For our studies we are talking about 140 studies spread out over about 200 collaborators. We have 40 sites worldwide. We collaborate with academic

institutions, biopharmaceuticals like GSK. We have been doing this for about -- we are probably the oldest and the largest academic DNA bank repository in the world.

I think the key for us, the difficulty from my end because I come from the research end and the biological end, is that we had to integrate all the lab programmers and the data management end but once you get that integration in and work together it is much easier but you definitely need a very strong IT database management personnel.

DR. RASOOLY: That actually was a concern in terms of the data repository. Is it feasible to hire people who will be able to get sufficient familiarity with several different data sets? We are not talking about hundreds but we are certainly talking about a dozen or more data sets. Or is there too much time investment involved in sort of assimilating the data? In other words, is it feasible to have somebody that has the 12 or 15 different CDs and who gets an inquiry, we need the 40 year old patients with diabetic nephropathy for six years who -- I do not know, whatever, and will a person be able to pick up that CD and find those patients in a particular study? Or is that an unrealistic expectation for a contractor?

DR. _____: Not unrealistic at all. In

fact, in the realm of the business world, our data sets are not that large at all. It is very manageable. What the standard is for both academics and the business world are called laboratory information management systems of which there is probably half a dozen or more commercially available third party software that are used by some of the other major corporations and universities such as ourselves.

DR. RASOOLY: Is being able to do research on these data sets a major interest of groups that would manage this? That is something that we have discussed over and over again. That might be something that would make a person more or less interested in managing these data sets. Everybody is nodding their head. Okay.

DR. _____: Well, it certainly makes, you know, the research data available to everybody is what essentially you envision down the road, correct?

DR. RASOOLY: Right.

DR. _____: You want these subsets available to all the basic researchers and that is easily manageable as well.

DR. _____: Just as a point of information, one of the things we built into all of our previous contracts is that the samples are available to researchers at Rutgers without any

charges.

DR. RASOOLY: I must have missed that one.

DR. _____: We did not.

(Laughter.)

DR. _____: Is that just within Rutgers itself?

DR. _____: A consortium of Rutgers, Robert Wood Johnson, UMDJS.

DR. _____: Okay.

DR. RASOOLY: I wanted to turn my attention briefly to the genetics part of the contract. I felt, and I have not heard anybody disagree, that the genetics is a unique aspect, that it is a unique set of tasks and not similar to any of the others, and one thing that I wanted to ask was a lot of our studies are considering the possibility of using frozen blood cells rather than making the transformed cell lines because whether the study is actually going to be a genetic study or not has not been determined. It is not, at first blush, a genetic study.

And the question is does the genetics contractor logically handle that task as well, receiving the sample, and cryopreserving the relevant cells or does that belong to the archival repository?

DR. _____: I think that would belong to the genetics contractor.

DR. RASOOLY: Is that tricky to do that properly, to cryopreserve?

DR. _____: No, not at all. No, but then in the event that -- when you are establishing the cell line you always have the back up of the frozen samples so it permits guaranteed generation of a cell line.

DR. RASOOLY: The reason I am asking that question is that could be thousands of samples that never become cell lines in the end and so it just basically would be archival storage then of material that is not too useful to anybody.

DR. _____: Well, it could serve as a source of DNA at some point.

DR. RASOOLY: Questions from -- I see there are some colleagues from other NIH institutes here. Are there any questions? Not to put you on the spot.

DR. _____: How large a back up blood repository are you envisioning?

DR. RASOOLY: How large a what?

DR. _____: A back up blood repository where you store bloods?

DR. RASOOLY: You know, we have had two disasters at NIH in this last six months or so. One was Hurricane Alison and the other was, of course, the events of September 11th, neither of which I

might add specifically, of course, affected NIH but we had researchers, in Texas who saw a life's worth of work wiped out and one could easily imagine the kind of catastrophic events of September 11th sort of doing a bad thing to a repository as well. And that has made us, at least made me, and I think I have persuaded my NIDDK colleagues, 100 percent committed to having a remote back up facility. I do not really see any alternative, frankly.

Now again it is going to be a cost issue. Are we going to back everything up? Are we going to back up only cell lines and DNA? What are we going to back up? But I think certainly for the cell lines and DNA we are going to insist on that. And in terms of the other samples, again it will be a cost issue.

For the data that is -- you know, it is not even an issue. Obviously things have to be backed up and backed up in some kind of remote facility so they can be accessed but that is a much less costly kind of operation.

DR. _____: By a "remote facility," is there a mileage distance between facilities that you are looking at?

DR. RASOOLY: I was not smart enough to figure out what that should be. I mean, I do not really know. I mean, I think if it is next door that the technical reviewers will have to evaluate whether

they consider that to be really a back up facility or whether it is the room next door or in the building on the way or what have you.

DR. _____: We have about 15 to 20,000 back up blood tubes now on campus and we divide that into two separate buildings and they are all on different -- they are both on different back up emergency systems, and we have not had a problem with that.

DR. _____: I think the issue with back up, as much as it needs to be physically separate to some extent, is the extent to which the back up can be managed and supervised. A remote back up that does not have physical presence constantly and is not monitored is not very useful.

DR. _____: No, it has to be monitored physically and electronically.

DR. _____: That is right.

DR. _____: 24/7.

DR. RASOOLY: At least one option for many of the ongoing studies is that they may just keep aliquots because they need aliquots of the samples anyway for what they are doing and send one aliquot to a repository, in which case de facto we have a back up facility and we have the primary study. So I think each study is going to be a little bit different and each sample is going to be a little bit

different.

DR. _____: Can I go in a different direction again? One of the assumptions you are going to make about the repository is that your intramural and extramural scientists will, in fact, submit the specimens. To what extent will you assure that that will happen and how will you assure that scientists, in fact, will comply with the submission requirements?

DR. RASOOLY: Which if one knew the answer to that question, one could become the director of NIH with no trouble.

(Laughter.)

I think that our hope, and I will defer to Dr. Hammond [NIDDK] after I finish, our hope is that we will for new studies make that a condition of award, that it will be clear from the outset that the samples will, in fact, be the property of NIDDK and deposited at a NIDDK repository at the conclusion of the study.

For studies that were not awarded under such conditions, what was it that they said? We depend on the kindness of strangers. We will have to have the PIs agree to that and, of course, the subjects will have to be consented on that.

Did you want to add something?

DR. HAMMOND: I agree with all those points.

Many of these upcoming applications would be submitted through RFAs, requests for applications. We would have that actually in the RFA. Some large studies which come in as unsolicited applications, we work with those in the notice of grant award to make sure we have that condition.

DR. _____: Let me follow up the question. One of the things that we know limits the scientist's intention to submit a specimen is they are still waiting for some publication or to complete their work. If they even, you know, send their specimens into the repository, would you also extend to them the courtesy that that would not -- materials would not be distributed further until they completed their publication and/or completed their research findings?

DR. RASOOLY: So, actually I think Jay [Everhart], this is more your question because what is really critical is not so much the samples. It is what fraction of the associated data will make those samples useful that somebody will actually want to analyze them. This is something that Dr. Everhart is really quite specialized in.

DR. EVERHART: Yes. I guess there are two parts to it. In most studies while the study is ongoing it is -- there is no release of data. I mean, you are collecting the data. It is not

finalized and it is the investigators who are using that.

At some point a study is actually completed and the investigators within that study have had ample opportunity to pursue their ideas within that study for those study materials, and at that point it would be opened up to a wider community, and it would be something that -- and exactly when that would be I think would be study dependent but there would be ample opportunity for the study investigators to use those materials and continue to use those materials but perhaps in a broader context.

The question of access to data I think is a little bit different and in terms of completed data sets, what would be on them, -- you know, the key thing is not to allow individuals to be identified on data sets but still have them to be robust enough to be used for linking to the appropriate samples and doing data analyses.

I guess one possibility actually with the data sets to make them a little bit easier to use is there could be our contractor who has essentially the complete data set and does not let it out and then there could be a very stripped down version of the data set that only contains some kind of key demographic and outcome data that would be widely distributed so anyone could look at the data set and

say, "Oh, yeah, this -- I might really be able to use this to answer a question and then pursue that." But that would be set up by study I would think, study by study. Does that address what you were talking about?

DR. _____: You know, for many large studies the study group through its publications committee decides on a registry of papers that the study will produce. Now some of those studies are, you know, Class E or F papers that might not be written for five years. And would the data related to those papers that were initiated by the study group all be protected so that only the study group could access those data or, in fact, you know, once the study was over and, you know, the core group of papers were written, would it essentially be open to anyone?

DR. EVERHART: This becomes kind of a very detailed technical point that is addressed down the line, I think, in our -- as Dr. Hammond mentioned in our notice of awards, we would be saying that, you know, ultimately these data are going to be used by a wider community and the study group has to be aware of that.

DR. _____: Information systems could be built such that you could have restrictions on who gets to see the data when and that there could be,

you know, upon consensus of the group to release at this point a release at a further date or set a time so that when the consortium gets together or the group that is doing the study actually sets a release information -- the information system could then make the data available.

So I think in building an information system to accommodate this, it is something to keep in mind that it could -- and I think the other thing to keep in mind is that it is a lot of effort to hold back depositing data till the end. A lot of people do find that as you are collecting the data it is easier to put it into a system as you are going along rather than waiting until the end and getting it all in at once.

DR. EVERHART: Yes.

DR. _____: That is a big task, especially as you have high rate of turnover in post-docs and/or graduate students and they are gone and pieces might be missing and it is very difficult to dig them up versus if they had an information system available to deposit it when they did the work right there and the fact that that data could be protected and/or have privileges assigned to it such that only certain people are able to see that data.

DR. RASOOLY: I think that is a really important point and one thing that Dr. Everhart has

emphasized is that one of the jobs of the data contractor is going to be to work with the data coordinating center before they stop existing to make sure that the data are in a form that is useable, that there is a manual that explains what each of the points are and how it was built and so on so that there will be a transitional process before the data are, so to speak, archived and that process could be quite long.

You are suggesting that it be even early on in the study as they are winding down the subject collection and moving into the -- you know, sort of the monitoring phase of the study, perhaps that would be a good time to begin to assemble the data set, you know, with a contractor so that it could be stored even though it would not be available for release for several years.

DR. EVERHART: I think if this is looked at prospectively, if we are aware that this is a process that is going to take place, it actually can work rather well. Where we have had trouble, problems is that at the end of a study we say, "Oh, well, it would be great if we had a public use data set and let's do it," and then it becomes quite difficult.

DR. _____: But you had mentioned before that you are thinking the data coordinating center is the one that is really doing more of the

data scrub and actual error checking and completeness checking and that kind of thing. So whether you think that at that point they are already having an information system that is available to their members and users that they are coordinating, I guess that was one thing that is not clear to me of where does the bigger archive information system pick up and where the data coordinating system information systems responsibility's lie.

Well, you had mentioned the concept of you have a data collection for each experimental design and that obviously they have in mind what pieces of information they want, and at least from my interpretation right now that you are thinking that they are doing the completeness checking and that at some point they would then be ushering this up to more of an archive-like public information system.

DR. EVERHART: It is really the responsibility of the data coordinating center to do all the quality control aspects of -- or to be responsible for the quality control aspects of the data. In fact, not all data sets are going to be archived. Some studies essentially go on -- are perpetual studies. For example, certain multi-center clinical studies where the cohort that was established is considered so important to follow that essentially they continue to be followed long after

the initial study period and that is usually a responsibility of the data coordinating center to collect and coordinate those data.

So in that circumstance the tissue repository would actually be working with the existing data coordinating center because there would be no reason to have had, you know, another contract to archive the data set until that study actually really does end.

DR. RASOOLY: Our primary concern is with a situation in which the data coordinating center is kind of going out of business and it is in that situation where we still have the samples that we need to archive the data so that people can figure out what samples there are and what data are associated with them, and that is where we see the database contractor.

But again in order for the contractor to be able to assume the data set they have to have worked with the data coordinating center before it stops existing. Otherwise, they are just getting a meaningless bunch of junk basically. Not to put too fine a point on it.

DR. _____: In that connection it seems like the level of effort that the repository contractor would be required to invest will vary widely depending on the quality of the data scrubbing

and cleaning and formatting and standardization that the data centers do and to the extent that they do not do that that suggests one level of effort and probably fairly difficult to predict in advance I would think.

DR. RASOOLY: I think that that is right. I mean, you know, you should be aware that the kinds of studies we are looking at are relatively large studies with relatively, you know, high level data coordinating centers. I think the level of chaos is pretty low but, you know, one should never overestimate the quality of things. You know, you hope that the data will be in good shape but you are right, there will be variable amounts of effort and also the willingness of the data coordinating center to make a manual that is useful and to explain the different aspects will vary.

I mean, that is just a personality issue and how easy it is to work with people and so on. You know, in that respect we do see that the contract as being variable over time as well as over cost but there will be some years and some periods of time that there will be much more effort than others.

DR. EVERHART: Again if this is looking prospectively and as we start studies, we are aware that ultimately the data is going to be moved from the data coordinating center to a third party that --

and so the data coordinating center has to be aware that that is going to happen and can set up their systems initially to do that so we know that, you know, they do not do everything on some idiosyncratic in-house database that cannot be, you know, transferred to something more robust and we need to -- we will make sure that is done in advance.

DR. _____: But if you are funding groups that already have a legacy system in place that could be challenging, especially if you are looking at trying to keep costs low that, you know, it is hard to get people to change if they already have invested heavily into pieces of hardware and software that they have already put in place.

DR. EVERHART: And just to be clear, we are talking about potentially many different data coordinating centers for future possible many different studies so a repository would potentially be working with many different types of databases.

DR. RASOOLY: Yes, absolutely. I mean, that is why we felt that we needed a person who was a real information systems expert that was flexible enough to learn the different systems and we are hoping that that is actually possible. I have had some reassurance here today and from others that it is possible and that the earlier on you start in the process, the better off you are.

On the other hand for us to spend a certain amount of money to preserve data that costs $15 million to collect, you know, -- the cost issue comes back again. If it costs us $15 million to preserve those data that is not worth it. If it will cost us, you know, $200,000 to preserve it, well, maybe that is worth it. So, you know, it always is a cost issue.

DR. GARFIELD: (Not at microphone.) One thing about the sort of ongoing studies, at least the multi-center studies, the actual processes tend to be fairly transparent. The data coordinating centers are involved. They are using standard platforms and software and they are constantly doing data analyses for the study and the -- our thinking is that for those larger studies that have that sort of transparency in ongoing data now it really would not be terribly difficult to move the data over. What is difficult is moving the expertise on manipulating those data over because personnel, of course, are often dedicated to the study know it better than anyone else and that is actually important.

DR. _____: Is it ever a consideration to put out a contract for a data coordinating center for one center to handle all potential clinical studies at NIDDK? Do you ever do something like that?

DR. RASOOLY: I think the feeling -- and I think my NIDDK colleagues will echo this -- is that sort of the thought of the study or the reason for the study, you know, resides with the ability to analyze the data and what makes people want to do these studies is their ability to sort of own and shape the research. So, generally when we put out RFAs for these very large studies that is a big attraction for people if they can run the data coordinating center and that brings in -- that sometimes is a nucleus for the groups that form to do this. So I am not -- I do not see us heading in that direction towards a central kind of analysis having all the, you know, clinicians out there collecting and us doing the analyzing.

DR. _____: Can I ask a qualifying question on that? I agree with what you just said that each study will have their own data coordinating center and each study will need to analyze their data but do you envision a time in the future that you would potentially offer a grant to a bioinformatics scientist who would do cross study data analysis and look for clues through data mining using this data set? And if you do -- the reason I am asking if you do, the design of the database becomes very critical to the success of that scientist who really needs to data mine to look for clues that might be very

invisible to an individual study section, individual grant but might become very visible if he has the opportunity to look at three, four or ten cross-sectional data sets.

DR. RASOOLY: Yes. That is certainly becoming an increasingly attractive idea. I think for purposes of the discussion today that is not our main intention. Right now, we are considering independent data sets that are not necessarily link-able or designed to do that. The most important thing is for that individual data set to be usable.

DR. RASOOLY: Would you like to talk a little bit about, I guess, sort of the next steps that we are going to do here?

MR. SULLIVAN: Yes. What the next phase we will be doing is reviewing everything that we have learned from you today. We will develop requests for proposals as a result of the meeting today. There is an internal process in the federal government where we do acquisition planning. We have to develop an acquisition plan, get the funds certified, get the funds available. Once that planning process is done then we release requests for proposals and we are hoping to have those RFPs out early next calendar year.

DR. RASOOLY: I wanted to ask Dr. Hammond to just say a few words about review and how we are

going to conduct the review.

DR. HAMMOND: Sometime down the road when the proposals actually arrive, they will arrive in the contracting office and that is Mr. Sullivan's area, and as he mentioned the number of these will determine not only the process for selection but the process for review in terms of how long it takes and what sort of panel we need to get together.

The way it works is that when the applications then arrive in our review branch, which is part of the Division of Extramural Activities, we have to make sure we have no conflicts between the reviewers selected and the offerors so we go through all the proposals very carefully because each reviewer who is on the technical evaluation panel scores each proposal. We do not have people leave the room like we can in grants. We have to have a group that is in no conflict at all.

This takes some time and then we would assemble the panel as I mentioned to go through these and since we do not know whether we will have a single RFP or multiple, we do not know exactly how this is going to work yet, but the overall time frame is it takes about three to four months from receipt of proposals in our contracting office to completion of the review and the reports which go to our contracting office to make sure we have the reviewers

have enough time to review the materials before the meeting and also to make sure we select the best panel.

All of the information then back and forth would go through our contracting officer. Unlike a grant, it does not come right to the review branch, you do not speak to our staff directly. If there are any questions about this once the proposal is submitted, the contracting office is the official contact point.

But in many ways it is like grant review. Of course, because we want the highest quality review of this, we want the best technical review, upon which we can finally make the final selections.

If you have any questions, in general, about the process I would be happy to answer them.

Okay. Thank you.

DR. RASOOLY: We were thinking that we would try to aim -- to give people a 90 day period to reply to these RFPs. That would be our target. If people feel strongly that that is too short or, well, too long -- no -- then, you know, this would be a good time to offer suggestions in that area. Is that an adequate period of time to prepare a response? It appears to be.

Well it looks as if we are done for this morning. We are going to be here all of us probably

till 12:00 o'clock if you would like to stay and talk to us a little bit more and I would ask in response to this meeting again if you have any comments please send them to the repository's mailbox within the next two weeks or so as we are beginning to do our planning. It would be very helpful.

Every comment you have sent so far has been extremely helpful and we have circulated it and read it and I assure you we give it the most serious attention.

DR. _____: Obviously we are recording this. How will that be available to parties that were not able to attend today?

DR. RASOOLY: It is being transcribed. There are tapes now and the tapes will be transcribed, and we are hoping to put it up in the same Q&A section of the web page where we have questions and answers now.

DR. _____: Rebekah [Rasooly], you may have covered this before I got in. I came in a few minutes late but, I mean, the scope of the studies that we are talking about are those that relate to the large multi-site clinical studies.

DR. RASOOLY: That is right.

DR. _____: And there are, you know, thousands of human studies out there that would not be considered for this.

DR. RASOOLY: That is right. That is right. This is going to be for large studies. We are not looking for the individual R01 investigator to send us, their 300 or 400 vials at all.

DR. _____: I was out of the room for this but you did address sort of the immortalization issue, the cell immortalization?

DR. RASOOLY: Yes. I mean, we talked about that briefly.

DR. _____: About what the reasonable retrieval rate is?

DR. RASOOLY: Success rate.

DR. _____: Success rate.

DR. RASOOLY: I have had some discussions. There are some people here and other places who are quite expert in this area. I have had some discussions and I am planning to write it for 97 percent and I see some people cringing but, I think if people think that is hopelessly unrealistic it might be worthwhile to write that to me and we will take that into consideration.

Obviously if Fed Ex's truck gets hijacked or something, I mean that is something else but within reason.

DR. _____: My only comment on that is I would separate that into domestic and international samples.

DR. RASOOLY: Absolutely.

DR. _____: We have collected in some very odd places in Micronesia and all sorts of places, and it takes a little bit longer.

DR. RASOOLY: Yes. Other questions?

DR. _____: I think another concern as I was at airport this morning, the post office is going to begin irradiation of all mail and all samples, and that is an issue that concerns us since we receive cell -- blood -- whole blood worldwide as well. That is another aspect of the study you are going to have to take into account. I think they use x-ray radiation.

DR. _____: (Not at microphone.) I am asking if you use the postal service.

DR. _____: We try not to. We found that at least domestically it takes a lot longer than we should reasonably expect but there are those odd ones. It is more than just the odd time but even within the State of North Carolina in the rural areas we have to send through normal U.S. Post.

One question I had was I was looking at your proposal and you are looking at 20,000 cell lines transformed per year?

DR. RASOOLY: Right.

DR. _____: That is really ambitious.

DR. RASOOLY: So that the number that was

contemplated.

DR. _____: Okay.

DR. RASOOLY: Was contemplated. That is the key word there. Over the next three to four years from NIDDK funded studies. And what I said at the outset, and I am sorry that apparently you were delayed at the airport and I hope they did not x-ray you --

(Laughter.)

that I collected these numbers the way you collect numbers, which is I asked every -- we surveyed every single study, large study that the NIDDK has done, is doing or will be doing in the coming few years, and collected all those numbers. These numbers are the maximum possible that any repository would ever handle at any point.

When we actually write the proposal -- the RFP, we will make a much more realistic evaluation of how many samples exist and need to be acquired and how many samples we anticipate for coming years, and we will obviously make that into a range.

I cannot estimate cell lines now except within the order of magnitude that it will probably be somewhere between 200 and 2,000 a year but that is just a ball park and that number may change. I do not want to commit myself to a number.

DR. _____: That is certainly

manageable.

DR. RASOOLY: But that is, I think, what we are looking at. No, not 20,000 in a year. We cannot collect patients that fast.

DR. _____: Rebekah [Rasooly], I have another follow-up question. I am sorry. Prior to the award do you expect to do a site visit?

DR. RASOOLY: Pat [Sullivan]?

MR. SULLIVAN: We will put in the RFP that it is our plan to make site visits and if that changes we will notify all offerors in the RFP.

DR. RASOOLY: Other questions? Thank you very much.

(Whereupon, the proceedings were concluded.)

* * * * *

1

2

3

4

5